# Measuring Patient-Perceived Quality of Care in U.S. Hospitals Using Twitter

## INTRODUCTION

Over the past decade, patient experiences have drawn increasing interest, highlighting the importance of incorporating patients' needs and perspectives into care delivery.[1 2] With health care becoming more patient centered and outcome and value driven, healthcare stakeholders need to be able to measure, report and improve outcomes that are meaningful to patients.[2-5] These outcomes can only be provided by patients, and thus systems are needed to be able to capture patient reported outcomes and facilitate the use of these data both on an individual patient level as well as on a population level.[2 3] Structured patient experience surveys such as Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) are common methods designed to assess patients' perception of the quality of their own health care.[4 6 7] A major drawback with these surveys is the signficiant time-lag - often several months - before official data is released, making it difficult for patients and other concerned stakeholders to be informed about current opinions on the quality of a given institution. Moreover, traditionally these surveys have low response rates, raising concerns about potential response and selection bias in the results.

Social media usage is pervasive in the U.S., with most networks seeing growth in their user base each year. As of 2014, approximately 1 out of 5 adults actively use Twitter; while most popular with adults under 50 years old, the network has seen significant growth in the 65 and older population in the past year.[8] Although there are legitimate privacy, social, ethical and legal concerns to interacting with patients on social media[9-11], it is clear that patients are using these venues to provide feedback.[12-20] Additionally, the use of social media data for health research has been gaining popularity in recent years.[21 22]

Sentiment analysis of social media is useful for determining how people feel about products, events, people and services.[23-27] It is widely used in other industries, including political polling[28-31] and brand/reputation management.[32 33] Researchers have also been experimenting with sentiment analysis of social media for healthcare research.[12 13 15 16] Sentiment can be determined in several ways, with the goal being to classify the underlying emotional information as either positive or negative. This can be done either purely by human input or by an algorithm trained to complete this process based on a human-classified set of objects.

We seek to describe the use of Twitter as a novel, real-time supplementary data stream to identify and measure patient-perceived quality of care in U.S. hospitals. This approach has previously been used to examine patient care in the UK.[16] While there was no correlation between Twitter sentiment and other standardized measures of quality, the analysis provided useful insight for quality improvement. Our aims are to provide a current characterization of U.S. hospitals on Twitter, explore the unsolicited patient experience topics discussed by patients and determine if Twitter data is associated with quality of care, as compared to other established metrics.

## METHODS

### HOSPITAL TWITTER DATA

We compiled a list of Twitter handles for each hospital in the United States. The Oct 2012 – Sept 2013 HCAHPS report served as our hospital master list and included 4,679 hospitals. We used Amazon's Mechanical Turk (AMT) - an online tool that allows large, tedious jobs to be completed very quickly by harnessing the efforts of crowdsourced employees[34] - to identify a Twitter account for each hospital. Two AMT workers attempted to identify an account for each hospital, with any disagreements resolved by manual inspection (FG). We utilized the services of DataSift, a data broker for historical Twitter data, to obtain all tweets that mentioned any of these hospitals during the 1-year period from Oct 1, 2012 - Sept 30, 2013. Mentions were defined as tweets that were specifically directed toward a hospitals' Twitter account (i.e., they included the full hospital Twitter handle, such as '@BostonChildrens'). During this time frame, we found 404,065 tweets that were directed at these hospitals. Tweets and associated metadata were cleaned and processed by custom Python scripts and stored in a database (MongoDB) for further analysis. This study only analyzed tweets that were completely public (i.e., no privacy settings were selected by the user) and that were original tweets – we ignored all re-tweets (tweets from another individual that have been re-posted) to ensure we were capturing unique patient experience feedback. Furthermore, there were no personal identifiers used in our analysis, and thus there was no knowledge of the users' identities. The study was approved by the Boston Children's Hospital Institutional Review Board and granted waiver of informed consent.

### MACHINE LEARNING CLASSIFIER

We manually curated a random subset of hospital tweets to identify those pertaining to patient experiences - defined as patient's, friend's, or family member's discussion of health care experience. Some examples of patient experience included: interactions with staff, treatment effectiveness, hospital environment (food, cleanliness, parking etc.), mistakes or errors in treatment or medication administration, and timing or access to treatment. Curation was achieved via two methods. The first method utilized a custom web-app that displayed random tweets from the database and allowed a curator (TR and KB) to label them if related to patient experiences. Each tweet was labelled by two curators, and only those tweets labelled identically were used for the analysis. The second method of data curation used AMT for crowd-sourced labelling. Again, multiple curators labelled each tweet and only those tweets that agreed in their labeling were used. After multiple rounds of curation, two curators rated 23,391 tweets using the web-app (agreement of 90.64%) and 15,000 tweets using AMT (agreement of 80.64%). These two sets were combined to create a training set of 2,216 tweets relating to patient experiences and 22,757 tweets covering other aspects of the hospital.

This training set was used to build a classifier that could automatically label the full database of tweets. The machine learning approach looks at features of the tweets (e.g., number of friends/followers/tweets from the user, user location and the specific words used in the tweet, but never username) and uses this information to develop a classifier. For the text of a tweet, we used a bag-of-words approach and included unigrams, bigrams and trigrams in the analysis. Specifically, we compared multiple different classifers (naive Bayes and support vector machine) and subjectively selected the best classifier based on a variety of metrics such as F1 score, precision, recall and accuracy. Building the classifier was an iterative process and we retrained and improved the classifier over many rounds of curation. We used 10-fold cross validation for evaluating the different classifiers, and selected a support vector machine classifier with an average accuracy of 0.895. This classifier on average, had an F1 score of 0.806, precision of 0.818, and recall of 0.795.

### SENTIMENT CALCULATION

We used natural language processing (NLP) to measure the sentiment of all patient experience tweets. Sentiment was determined using the open-source Python library TextBlob.[35] The sentiment analyzer implementation used by TextBlob is based on the Pattern library[36], which is trained from human annotated words commonly found in product reviews. Sentiment scores range from -1 to 1. Sentiment scores of exactly 0.0 were discarded, as they typically indicate that there were not sufficient data. The average number of patient experience tweets for all hospitals was 43. To ensure there were enough tweets to provide an accurate assessment of sentiment, we calculated a mean sentiment score for each hospital with ≥50 patient experience tweets (n=297).

## HOSPITAL CHARACTERIZATION

We compared the proportion of hospitals in each of the following American Hospital Association (AHA) categorical variables between the highest and lowest sentiment quartiles: region, urban status, bed count, nurse-to-patient ratio, profit status, teaching status and percent of patients on Medicare/Medicaid. We compared nurse-to-patient ratio and percent of patients on Medicare/Medicaid to the median national value. We used the following Twitter characteristics (measured in August 2014) for sentiment correlation and quartile comparison: days account has been active, number of status updates, number of followers, number of patient experience tweets received and number of total tweets received.

## TOPIC CLASSIFICATION

We again utilized AMT to identify what topics were being discussed in the patient experience tweets. A total of 11,602 machine identified patient experience tweets were classified by AMT workers as belonging to one or more predefined categories. Only tweets with agreed upon labels were further analyzed; this totalled 7,511 tweets (agreement of 64.7%). Of these, 3,878 were identified as belonging to a patient experience category, while 3,633 were found to not truly be about patient experience.

## HOSPITAL SURVEYS

We emailed contacts with formal positions in the office of patient or public relations (or equivalent) of the 297 hospitals with ≥50 patient experience tweets (111 unique Twitter accounts) and asked them to provide feedback regarding their use of Twitter for patient relations. If employees could not be identified, either the department email (n=44) or general contact email (n=40) was used. Contact was attempted twice, with a second email sent 9 days after the first if necessary. The questions asked were: 1) “Does your hospital monitor Twitter activity?”; 2) “Do you follow-up with patients regarding comments they make on Twitter?”; and 3) “Are you aware that patients post about their hospital/care experience on Twitter?”. Informants were told their participation in the study was voluntary, confidential and anonymous.

## COMPARISON TO VALIDATED MEASURES OF QUALITY OF CARE

We chose two validated measures of quality of care. The first was HCAHPS, the formal U.S. nationwide patient experience survey. The intent of the HCAHPS is to provide a standardized survey instrument and data collection methodology for measuring patients' perspectives on hospital care, which enables valid comparisons to be made across all hospitals. Like other traditional patient surveys, the HCAHPS is highly standardized and well-validated.[4 6 7] We focused on the percent of patients who rated a hospital a 9 or 10 (out of 10), which has been shown to correlate with direct measures of quality.[4] We analyzed data from the HCAHPS period of October 1, 2012 – September 30, 2013. The second was the Hospital Compare 30-day hospital readmission rate calculated from the period of July 1, 2012 – June 30, 2013. This is a standardized metric covering 30-day overall rate of unplanned readmission after discharge from the hospital and includes patients admitted for internal medicine, surgery/gynecology, cardiorespiratory, cardiovascular, and neurology services.[37] The score represents the ratio of predicted readmissions (within 30-days) to the number of expected readmissions, multiplied by the national observed rate.[38]

## STATISTICAL ANALYSES

We used Pearson’s correlation to assess the linear relationship between numeric variables, Fisher’s exact test to compare proportions between categorical variables and a two-tailed independent t-test to compare the means of quantiles. Bonferroni correction was used to adjust for multiple comparisons. Ordinary Least Squares multivariable linear regression was used to adjust for potential confounders. Potential confounders extracted from AHA included: region, size, bed count, profit status, rural/urban status, teaching status, nurse-to-patient ratio, percentage of patients on Medicare and percentage of patients on Medicaid. Twitter account confounders (total statuses, total followers and total days since account creation) were measured in August 2014. Additional Twitter covariates were the total number of patient experience tweets received during the study period and whether or not the hospital had a unique Twitter handle (as opposed to sharing with a larger healthcare network). A Wald test was used to test for trend signficance.

## RESULTS

### CHARACERISTICS OF U.S. HOSPITALS ON TWITTER

Of the 4,679 U.S. hospitals identified, 2,349 (50.2%) had an account on Twitter; this included data from 1,609 Twitter handles (as many hospitals within a provider network share the same Twitter handle). During the 1-year study period, we found 404,065 total tweets directed towards these hospitals (data from 1418 Twitter handles, representing 2137 hospitals); of these, 369,197 (91.4%) were original tweets (data from 1417 Twitter handles, representing 2136 hospitals). The classifer tagged 34,725 (9.4%) original tweets relating to patient experiences and 334,472 (90.6%) relating to other aspects of the hospital. Patient experience tweets were found for 1,065 Twitter handles, representing 1,726 hospitals (36.9%).

Table 1 describes the common characteristics for all of the hospitals on Twitter. Overall, the mean number of patient experience tweets received for all hospitals during the 1-year study period was 43. The median sentiment values for the highest and lowest quartiles were 0.362 and 0.211, respectively. The proportion of hospitals in the profit status ($p<0.001$) and bed count ($p=0.023$) categories was significantly different between the highest and lowest sentiment quartiles, with public and larger hospitals overrepresented in the lowest sentiment quartile.

**Table 1: Characteristics of US Hospitals using Twitter**

| | Followers (n=2349) | | Sentiment (n=297) | | Proportion of Hospitals in Sentiment Quartiles | | |
|---|---|---|---|---|---|---|---|
| **Metric** | **Median** | **IQR** | **Median** | **IQR** | **Highest Quartile** | **Lowest Quartile** | **p-value** |
| **Region** | | | | | | | 0.392 |
| Northeast | 666 | 188 – 2686 | 0.278 | 0.124 – 0.377 | 0.27 | 0.31 | |
| Midwest | 981 | 176 - 2881 | 0.296 | 0.263 – 0.332 | 0.33 | 0.20 | |
| West | 437 | 118 - 1426 | 0.213 | 0.213 – 0.293 | 0.02 | 0.17 | |
| South | 832 | 183 - 2522 | 0.300 | 0.280 – 0.334 | 0.39 | 0.31 | |
| **Urban** | | | | | | | 0.379 |
| Yes | 1087 | 303 – 3069 | 0.293 | 0.244 – 0.334 | 0.77 | 0.93 | |
| No | 364 | 70 - 1871 | 0.301 | 0.263 – 0.334 | 0.23 | 0.07 | |
| **Bed count** | | | | | | | 0.037* |
| Small (<100) | 439 | 72 – 2198 | 0.312 | 0.270 – 0.338 | 0.41 | 0.13 | |
| Medium (100-299) | 622 | 166 – 2182 | 0.294 | 0.270 – 0.334 | 0.24 | 0.30 | |
| Large (300+) | 1610 | 527 – 3592 | 0.280 | 0.222 – 0.331 | 0.35 | 0.57 | |
| **Nurse-patient ratio** | | | | | | | 0.395 |
| Above national | 853 | 151 – 3078 | 0.301 | 0.270 – 0.338 | 0.59 | 0.39 | |
| Below national | 741 | 182 - 2199 | 0.283 | 0.223 – 0.334 | 0.41 | 0.61 | |
| **Profit status** | | | | | | | <0.001** |
| Public | 237 | 48 – 1549 | 0.263 | 0.112 – 0.299 | 0.06 | 0.26 | |
| Private nonprofit | 1115 | 281 – 3008 | 0.301 | 0.263 – 0.334 | 0.88 | 0.74 | |
| Private for-profit | 327 | 103 – 934 | 0.280 | 0.281 – 0.326 | 0.06 | 0.00 | |
| **Teaching hospital** | | | | | | | 0.242 |
| Yes | 1359 | 382 – 3498 | 0.285 | 0.223 – 0.332 | 0.45 | 0.67 | |
| No | 527 | 119 - 2005 | 0.301 | 0.270 – 0.334 | 0.55 | 0.33 | |
| **Medicare** | | | | | | | 0.617 |
| Above national | 605 | 138 – 2185 | 0.298 | 0.265 – 0.334 | 0.55 | 0.37 | |
| Below national | 1187 | 247 - 3228 | 0.294 | 0.228 - 0.334 | 0.45 | 0.63 | |
| **Medicaid** | | | | | | | 1 |
| Above national | 756 | 160 – 2459 | 0.295 | 0.227 – 0.338 | 0.61 | 0.65 | |
| Below national | 818.5 | 187 - 2828 | 0.298 | 0.270 – 0.334 | 0.39 | 0.35 | |

We found no correlation between sentiment and Twitter characteristics, except a weak negative correlation (r=-0.18, p=0.002) with total days the account was active. When the highest and lowest sentiment quartiles were compared after ranking hospitals based on these characteristics, only the total number of tweets was shown to have an effect on sentiment (p=0.002).

Hospitals with 50+ tweets were more active on Twitter, as they had more posts and followers (p<0.001), but their account was not older. In addition, hospitals with more patient experience posts were more likely to be below the national median of Medicare patients (p<0.001), above the national median for nurse-patient ratio (p=0.006) and non-profit (p<0.001). Figure 1 shows the geographical distribution of all U.S. hospitals on Twitter, highlighting sentiment and number of patient experience tweets received.

TOPIC CLASSIFICATION

We identified the topics of patient experience that were discussed in a random subset of tweets (Table 2). Box 1 includes some specific examples of each topic.

**Table 2: Topic Classification**

| Topic | Count | Ratio of Pos/Neg Tweets | Sentiment Median |
|---|---|---|---|
| Discharge | 6 | 0.500 | -0.096 |
| Time | 313 | 0.514 | -0.150 |
| Side-effect | 10 | 0.667 | -0.150 |
| Communication | 205 | 0.884 | -0.039 |
| Money | 222 | 0.917 | -0.028 |
| Pain | 37 | 1.000 | -0.007 |
| Room Condition | 41 | 1.769 | 0.140 |
| Medication Instructions | 10 | 2.000 | 0.138 |
| Food | 35 | 2.625 | 0.250 |
| General | 2999 | 6.734 | 0.467 |
| **Totals** | 3878 | 3.762 | 0.400 |

Topics are ordered based on the ratio of positive to negative tweets.

**Box 1: Patient Experience Tweets**

- Discharge instructions (including care at home)
  - "…epic fail on my TKA discharge"
- Time management
  - "12 hrs in the ER..... not good [hospital]. My father is just now getting a room. #notgoodenough #er #getitright"
- Treatment side-effects
  - "Im on 325mg aspirin coming off blood thinners after blood clott found in lung will this affect my heart??"
- Communication with staff
  - "I know it's Monday/Flu season but waiting 3.5 hrs for pediatrics to call me back is a little much"
- Money concerns
  - "Hey [hospital], can you hold off on the collections calls until my bill is actually due? Please try to keep it classy."
- Pain management
  - "[hospital] pediatric ER sucks! No doctors to assist, nurses in the back having coffee while there is a sick child in pain in empty ER"
- Conditions of rooms/bathrooms
  - "what do you have against coat hooks? 3 exam rooms, 2 locations in 1 day and not 1 place to hang my coat and bag!"
- Medication instructions
  - "[hospital] gave my mom a prescription for a discontinued medicine. #Silly hospital. @Walgreens saved the day!"

- Cafeteria food
    - “skipped dinner last night because of your terrible cafeteria food. Are you trying to get more patients? #eathealthy #healthcare”
- General satisfaction/dissatisfaction with procedure and/or staff
    - “I'm thankful for the [hospital], their staff, my Doctors and to be treated as a person, not just a patient.”

USE OF TWITTER DATA BY U.S. HOSPITALS

49.5% of the 297 hospitals surveyed about Twitter use, responded. All hospitals indicated that they monitored Twitter closely, actively interacted with patients via Twitter, and were aware that patients post about their care experiences. Box 2 includes some additional representative feedback received.

**Box 2: Hospital responses regarding their use of patient Twitter posts**

- "[O]ur goal is to respond to patient comments within an hour of their post. If the comment can be addressed via Twitter, we direct them to appropriate resources online. We are extremely careful to abide by HIPAA guidelines and the protection of patient privacy."
- “We also do social listening to find out what topics are important to our patient families and supporters so we can take a proactive approach and participate in the conversation by providing information and help.”
- “We’ve had patients tweet us from waiting rooms, we’ve even had patients tweet us from their hospital beds!”
- “We proactively respond to all social media conversations that mention us. We respond then try to facilitate a one-on-one email or phone conversation with the patient or caregiver to discuss their experience.”
- "We utilize geocoding to narrow tweets to within .25 kilometers of each facility to capture any tweets that don’t expressly mention [hospital network name], but are related to their health care experience at a facility."
- “If there is a serious complaint, we send those along to Patient Relations who may follow up outside of social media. We typically don't get into a back and forth around a negative comment, but rather let the person know we are sorry for their experience, and then direct them to patient relations.”

LINKING TWITTER DATA TO QUALITY OF CARE

We observed a correlation between 30-day readmission rates and sentiment. There is a moderate negative correlation ($r=-0.215$, $p<0.001$), where higher sentiment scores are associated with lower 30-day readmission rates (Figure 2). Additionally, there was a significant difference ($p=0.014$) between the 30-day readmission rates in the highest vs. lowest quartiles of hospitals ranked on sentiment. Finally, after adjustment for hospital and Twitter characteristics using multivariate linear regression there was still a significant association between higher sentiment scores and lower 30-day readmission rates (Table 3; $p=0.003$).

**Table 3: Sentiment Associated with 30-day Readmission Rates**

| Mean Sentiment | 30-day Readmission Rate (Adjusted Score) | p-value |
|---|---|---|
| Lowest quartile | 16.876 | ref |
| Second quartile | 16.937 | 0.799 |
| Third quartile | 16.249 | 0.009 |
| Highest quartile | 16.163 | 0.009 |
| p-value for trend | 0.003 | |

In the univariate analysis, we found a significant difference between percent of people giving an HCAHPS rating of 9 or 10 for hospitals that have a Twitter account compared to those that do not (0.71 vs. 0.69, $p=0.001$) and between HCAHPS rating in the highest vs. lowest quartiles of hospitals ranked by the number of Twitter followers (0.72 vs. 0.69, $p<0.001$). Additionally, there was a significant difference in sentiment in the highest vs. lowest quartiles of hospitals ranked on HCAHPS (0.30 vs. 0.26, $p=0.017$). However, after adjusting for potential confounders through multivariate linear regression we did not find any significant correlation between HCAHPS and any of these metrics.

## DISCUSSION

Our findings indicate that patients use Twitter to provide feedback about the quality of the care they receive at U.S. hospitals. We found that approximately half of the hospitals in the U.S. have a presence on Twitter and that sentiment towards hospitals was, on average, positive. Of the 297 surveyed, half responded and all confirmed that they closely monitor social media and interact with users. We therefore conclude that the stakeholders of these hospitals see the value of capturing information on the quality of care in general, and patient experience in particular. We also found associations between Twitter sentiment and an established objective measure of hospital quality. Taken together, our findings suggest that Twitter is a unique platform to engage with patients and to collect feedback - and possibly a useful measure for supplementing traditional approaches of assessing and improving quality of care.

Our findings on the extent of Twitter usage by hospitals are similar to what has been reported previously.[39] The generally positive sentiment on Twitter is consistent with other analyses that suggest a positive language bias on social media.[40] However, our analysis of Twitter sentiment, and its association with conventional quality measures, is novel. There were some striking differences between the hospitals with the highest and lowest sentiment, with both large and public hospitals being over-represented in the lowest quartile. Additionally, the number of tweets a hospital received, in part, influenced hospital sentiment; hospitals that received more tweets had, on average, higher sentiment. However, the number of tweets a hospital posted did not affect sentiment. Thus, having a more active online presence with a frequent posting behavior is not sufficient to increase sentiment alone, though we did find that it increased the likelihood of receiving more patient experience tweets.

Twitter feedback is entirely unsolicited. As such, there was a wide range of patient experience topics discussed. These topics include those covered by the HCAHPS survey, as well as some not addressed (e.g., time, side-effects, money and food concerns). It is not surprising that some topics tended to be more negative than others - for example, discussion of time, money or pain is not likely to be positive. Thus, from an individual hospital's perspective, it might not be useful to heavily weight the number of positive or negative tweets within one topic category at any one moment. However, monitoring these topics over time and detecting when sentiment goes above or below an established baseline could be useful.

We used 30-day hospital readmission rates as an objective measure of quality of care. This metric has been used before, including recent studies that showed correlation with ratings on Facebook[14] and Yelp[17]. Similar to these studies, we found a moderate negative correlation between Twitter sentiment and readmission, with higher sentiment hospitals having a lower readmission rate. This association survived adjustment of potential confounders, with a significant downward trend for readmission rates as sentiment increases.

Additionally, we assessed the percent of patients who rated a hospital highly on the HCAHPS survey. While there were some interesting associations, these were not significant in the multivariate analyses. One caveat to this comparison is that HCAHPS does not ask specifically ask questions about a friend or family member's experience receiving care.

There were several limitations to our analysis. We only looked at tweets that explicitly included a hospital's Twitter handle. Broadening our criteria to include hospital names as keywords or attempting to assign tweets to nearby hospitals given geospatial data, could have potentially increased the number of patient experience tweets we identified. Additionally, many hospitals within a larger network shared a Twitter account and, without additional follow-up, it is difficult to determine which hospital is being discussed. Like all surveys, our hospital questionnaire may have been subject to a potential response and selection bias. Due to the cross-sectional design, while we have shown association between organizations that use Twitter and their interactions with patients, we cannot confirm any causal relationship. Further investigation of how these findings change over time would be helpful. Finally, while patient experience classification had relatively high agreement rates (over 80%), topic classificaiton only had an agreement of 64.7%. This is likely an effect of using crowdsourced curators without a high level of domain-specific training, which also explains why 77.3% of patient experience tweets were non-specifically labelled as 'General'. As for our automated approach, machine classification and sentiment analysis using NLP does not perform as well as human curation. With these caveats acknowledged, our approach enabled processing of an extremely large amount of data and illustrated that automated analysis of Twitter data can provide useful, unsolicited information to hospitals across a wide variety of patient experience topics.

Our findings have implications for various groups. Hospital administrators and clinicians should actively monitor what their patients are saying on social media. Institutions that don't use Twitter should create accounts and analyze the data, while existing users should consider leveraging automated tools. Insight from key leaders at institutions will help to better understand gaps and potential opportunities. Regulators should consider social media commentary as a supplemental source of data about care quality. The information is plentiful and, although the techniques for processing and understanding this data are still being developed and improved, potentially important. We recommend a larger survey in the U.S. and globally with all relevant stakeholders, including patients and their families, to get a better understanding of the use and value of social media for patient interactions. The public should pay attention to what other people are tweeting - and systems to collect, aggregate and summarize these tweets for a public audience in real-time should be considered to complement data from traditional reporting platforms. To increase the utility of this data, we would recommend that each hospital manages their own unique Twitter identity, rather than share an account across a larger healthcare network.

## CONCLUSIONS

We show that monitoring Twitter provides useful, unsolicited, and real-time data that might not be captured by traditional feedback mechanisms. These data correlate with a validated measure of quality of care. While many hospitals monitor their own Twitter feeds, we recommend that patients, researchers and policy makers also attempt to utilize this data stream to understand the experiences of healthcare consumers, and the quality of care they receive.


## ACKNOWLEDGMENTS

We would like to thank the hospitals that graciously participated in our survey. We would also like to thank Clark Freifeld for his help in acquiring the raw data from DataSift.


## CONTRIBUTORSHIP

JBH had full access to all of the raw data in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. Study design: JBH, JSB, and FG. Acquisition of data: JBH and CF. Manual Curation: TR and KB. Machine learning: JBH, GT. Analysis and interpretation of data: JBH, JSB, TR, KB, EON, DJM, RR, AW, FTB, FG. Drafting of manuscript: JBH and FG. Statistical analysis: JBH, TR, KB and DJM. Critical revision of the manuscript for important intellectual content: JBH, JSB, TR, KB, EON, DJM, RR, AW, FTB and FG. Study supervision: JSB and FG.

## COMPETING INTERESTS

The authors have no competing interests to declare.

## DATA SHARING STATEMENT

Aggregate data may be obtained by writing to the corresponding author.


## FUNDING

Funded by NLM Training Grant T15LM007092 (JBH), NLM 1R01LM010812-04 (JSB), and a Harkness Fellowship from the Commonwealth Fund (FG).

**REFERENCES**


1. Committee on Quality of Health Care in America, Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington (DC), 2001.
2. Rozenblum R., Miller P., Pearson D., et al. Patient-centered healthcare, patient engagement and health information technology: the perfect storm. In: Grando MA, Rozenblum R, Bates DW, eds. Information technology for patient empowerment in healthcare. 1st ed. Berlin: Walter de Gruyter GmbH & Co. KG, 2015:p. 3-22.
3. Zimlichman E. Using patient reported outcomes to drive patient-centered care. In: Grando MA, Rozenblum R, Bates DW, eds. Information technology for patient empowerment in healthcare. 1st ed. Berlin: Walter de Gruyter GmbH & Co. KG, 2015:p. 241-56.
4. Jha AK, Orav EJ, Zheng J, et al. Patients' perception of hospital care in the United States. N Engl J Med 2008;**359**(18):1921-31.
5. US Government. The Patient Protection and Affordable Care Act. 2009.
6. Goldstein E, Farquhar M, Crofton C, et al. Measuring hospital care from the patients' perspective: an overview of the CAHPS Hospital Survey development process. Health Serv Res 2005;**40**(6 Pt 2):1977-95.
7. Darby C, Hays RD, Kletke P. Development and evaluation of the CAHPS hospital survey. Health Serv Res 2005;**40**(6 Pt 2):1973-6.
8. Maeve Duggan, Nicole B. Ellison, Cliff Lampe, et al. Social Media Update 2014. January 9, 2015 2015. http://www.pewinternet.org/2015/01/09/social-media-update-2014/.
9. Hawn C. Take Two Aspirin And Tweet Me In The Morning: How Twitter, Facebook, And Other Social Media Are Reshaping Health Care. Health Affairs 2009;**28**(2):361-68.
10. Hader AL, Brown ED. Patient privacy and social media. AANA J 2010;**78**(4):270-4.
11. Moses RE, McNeese LG, Feld LD, et al. Social media in the health-care setting: benefits but also a minefield of compliance and other legal issues. Am J Gastroenterol 2014;**109**(8):1128-32.
12. Huesch MD, Currid-Halkett E, Doctor JN. Public hospital quality report awareness: evidence from National and Californian Internet searches and social media mentions, 2012. BMJ open 2014;**4**(3):e004417.
13. Wong CA, Sap M, Schwartz A, et al. Twitter sentiment predicts Affordable Care Act marketplace enrollment. J Med Internet Res 2015;**17**(2):e51.
14. Glover M, Khalilzadeh O, Choy G, et al. Hospital Evaluations by Social Media: A Comparative Analysis of Facebook Ratings among Performance Outliers. J Gen Intern Med 2015.
15. Wallace BC, Paul MJ, Sarkar U, et al. A large-scale quantitative analysis of latent factors and sentiment in online doctor reviews. Journal of the American Medical Informatics Association : JAMIA 2014;**21**(6):1098-103.
16. Greaves F, Laverty AA, Cano DR, et al. Tweets about hospital quality: a mixed methods study. BMJ quality & safety 2014;**23**(10):838-46.
17. Verhoef LM, Van de Belt TH, Engelen LJ, et al. Social media and rating sites as tools to understanding quality of care: a scoping review. J Med Internet Res 2014;**16**(2):e56.
18. McCaughey D, Baumgardner C, Gaudes A, et al. Best Practices in Social Media: Utilizing a Value Matrix to Assess Social Media's Impact on Health Care. Social Science Computer Review 2014;**32**(5):575-89.
19. Timian A, Rupcic S, Kachnowski S, et al. Do patients "like" good care? measuring hospital quality via Facebook. Am J Med Qual 2013;**28**(5):374-82.
20. Rozenblum R, Bates DW. Patient-centred healthcare, social media and the internet: the perfect storm? BMJ quality & safety 2013;**22**(3):183-6.
21. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection--harnessing the Web for public health surveillance. N Engl J Med 2009;**360**(21):2153-5, 57.
22. You are what you Tweet: Analyzing Twitter for public health. ICWSM; 2011.
23. S A, BA H. Predicting the Future with Social Media. IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT) 2010:492–9.
24. Coviello L, Sohn Y, Kramer AD, et al. Detecting emotional contagion in massive social networks. PloS one 2014;**9**(3):e90315.
25. Golder SA, Macy MW. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. Science 2011;**333**(6051):1878-81.
26. Haney C. Sentiment Analysis: Providing Categorical Insight into Unstructured Textual Data.
27. Pang B, Lee L. Opinion mining and sentiment analysis. Foundations and trends in information retrieval 2008;**2**(1-2):1-135.

28. A Preliminary Investigation into Sentiment Analysis of Informal Political Discourse. AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs; 2006.
29. A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. Proceedings of the ACL 2012 System Demonstrations; 2012. Association for Computational Linguistics.
30. Tumasjan A, Sprenger TO, Sandner PG, et al. Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. ICWSM 2010;**10**:178-85.
31. Ceron A, Curini, L., Iacus, S. M., & Porro, G. Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. New Media & Society 2014;**16**(2):340-58.
32. Jansen BJ, Zhang M, Sobel K, et al. Twitter power: Tweets as electronic word of mouth. Journal of the American society for information science and technology 2009;**60**(11):2169-88.
33. Ghiassi M, Skinner J, Zimbra D. Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. Expert Systems with applications 2013;**40**(16):6266-82.
34. Peer E, Vosgerau J, Acquisti A. Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. Behavior research methods 2014;**46**(4):1023-31.
35. Loria S. TextBlob: Simplified Text Processing. 2014. https://textblob.readthedocs.org/en/dev/index.html.
36. Center) CCLPR. pattern.en Documentation. 2015. http://www.clips.ua.ac.be/pages/pattern-en.
37. Medicare.gov. 30-day unplanned readmission and death measures. 2015. http://www.medicare.gov/hospitalcompare/Data/30-day-measures.html?AspxAutoDetectCookieSupport=1.
38. Centers for Medicare and Medicaid Services. Frequently Asked Questions: CMS Publicly Reported Risk-Standardized Outcome and Payment Measures. 2015. http://www.qualitynet.org/dcs/ContentServer?cid=1228774724512&pagename=QnetPublic%2FPage%2FQnetTier4&c=Page.
39. Griffis MH, Kilaru SA, Werner MR, et al. Use of Social Media Across US Hospitals: Descriptive Analysis of Adoption and Utilization. J Med Internet Res 2014;**16**(11):e264.
40. Dodds PS, Clark EM, Desu S, et al. Human language reveals a universal positivity bias. Proc Natl Acad Sci U S A 2015;**112**(8):2389-94.

## FIGURES

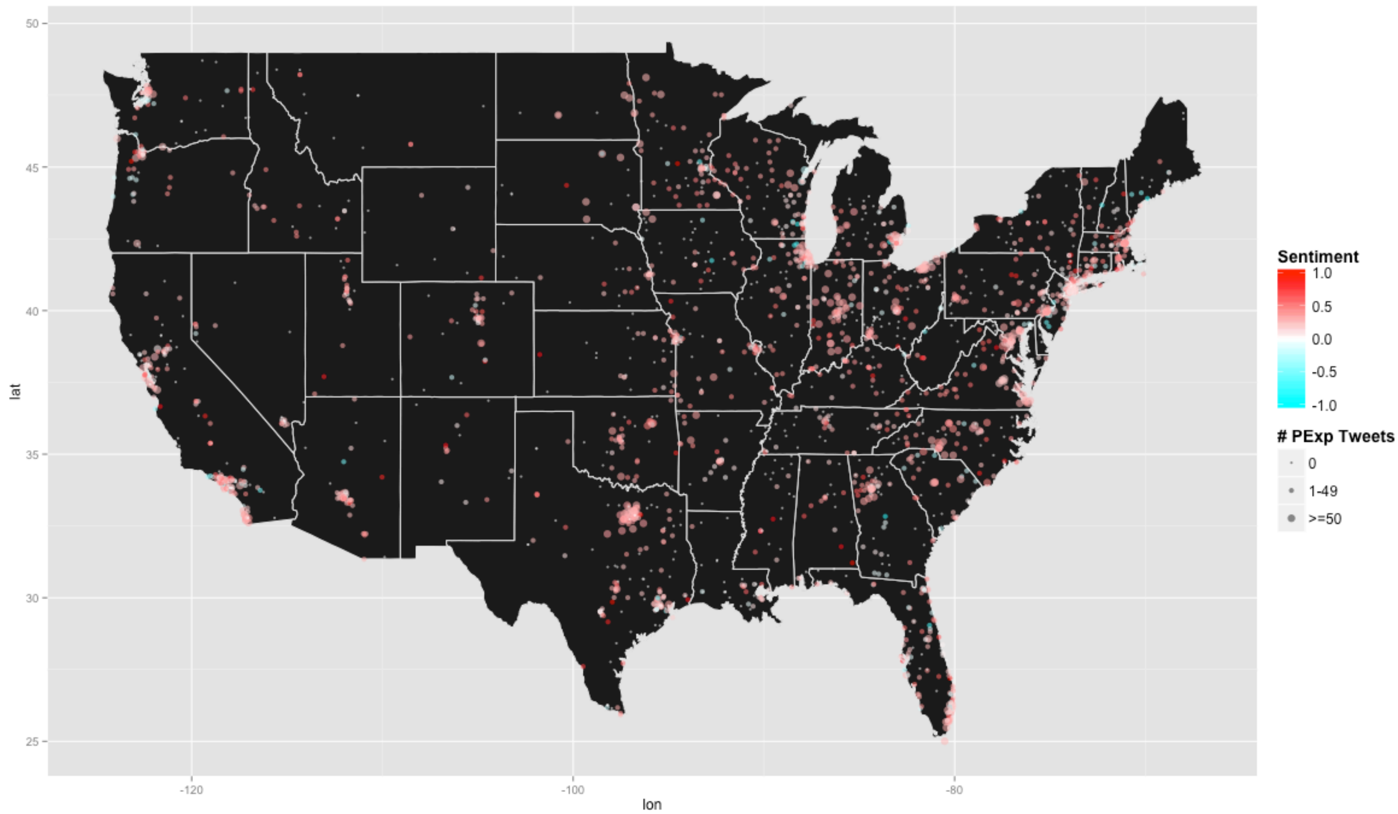


**Figure 1: Geographic Distribution of U.S. Hospitals on Twitter.**
Geographical distribution of all U.S. hospitals on Twitter (n=2,349). Hospitals are colored by mean sentiment, and sized by the number of patient experience tweets received in the 1-year study period. Sentiment ranges from -1 (negative) to 1 (positive).

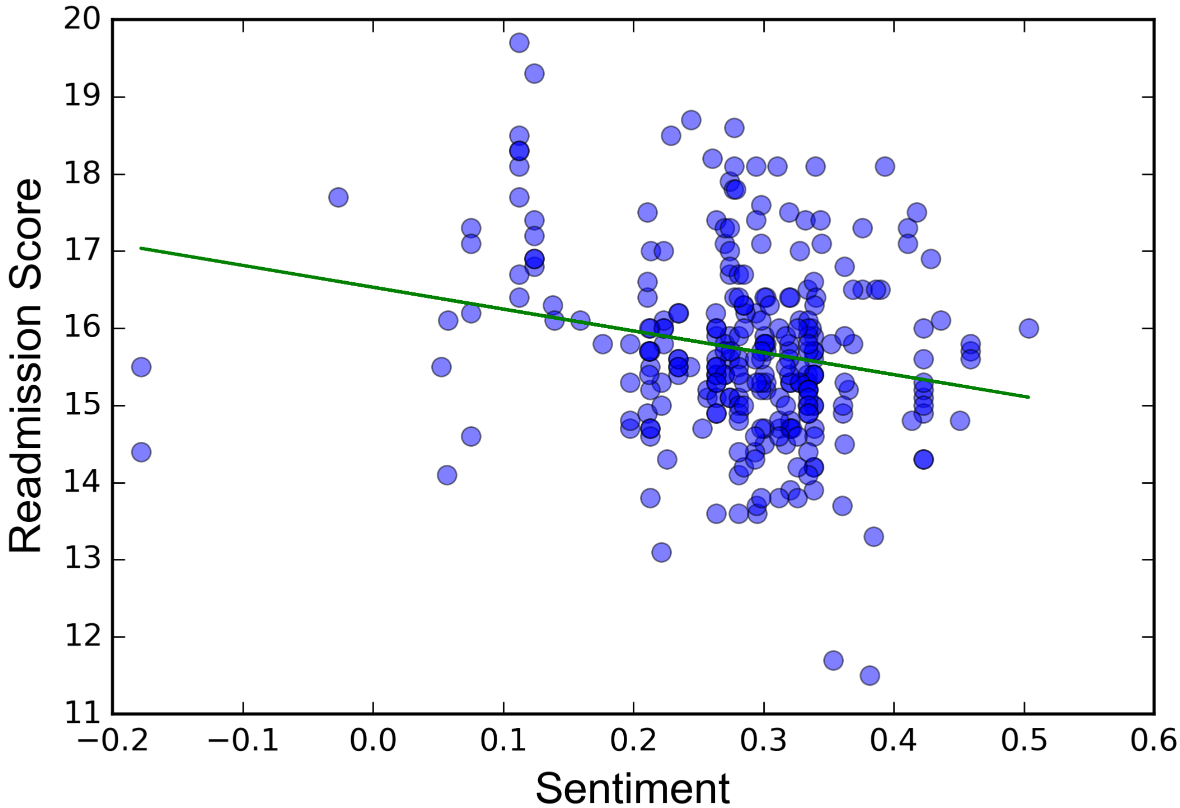


**Figure 2: Sentiment is Correlated with 30-day Hospital Readmission Rates.**
30-day hospital readmission rates are plotted against average sentiment, for hospitals that have ≥50 patient experience tweets (n=297). This association displays a moderate negative correlation (r=-0.215, p<0.001).